



## 方言研究の意義

伝統方言は死滅しつつある文化資源である。

北京語言大学『漢語方言地図集』

→金沢大学“日中無形文化遺産”連携融合  
事業とのコラボレーション

- 特徴
1. 中国初の項目別全国地図集
  2. 調査規模においてフランス（ALF）、ドイツ（Wenker Atlas）、日本（LAJ）等に匹敵する。

## なぜ中国語方言のデータベースをつくるか？

中国の研究においては言語地理学的な観点が欠如している。

＊言語地理学の目的

1. 方言地図から単語・言語特徴の歴史を再構成する。cf. 『蝸牛考』『全国アホバカ分布地図』
2. 言語変化の普遍性を解明する。  
譬えて言えば「がらくた遺伝子」の研究

## Project on Han Dialects (PHD)

- ▣ 1989～2006
- ▣ 共同研究(科研総合A3回, 基盤B2回)
  - ▣ 中国全土の方言調査記録に基づく
  - ▣ 方言データの集積とデータベース
  - ▣ 方言(解釈)地図の作成
  - ▣ 全国の研究者と連携

## PHDシステムの開発

- 2004-2006年科研基盤( B )による  
“新 P H D プロジェクト”
- P H D システム製作者：林智
- 金沢大学にサーバ・システムを設置

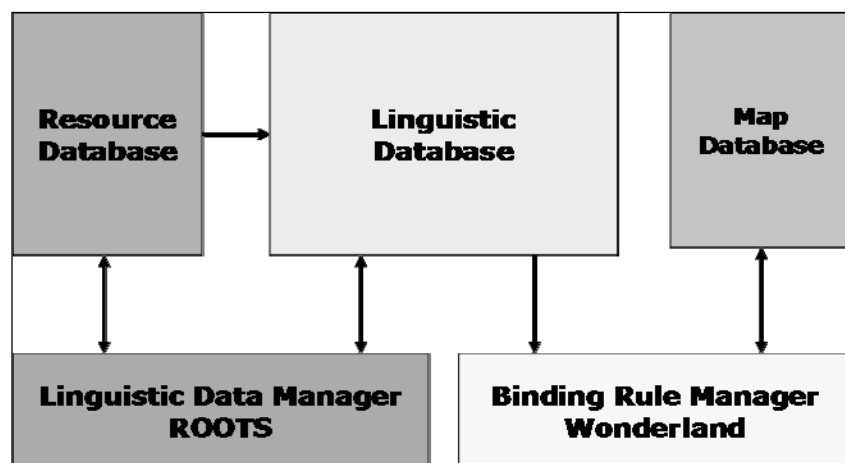
## システムの目的

- 言語データベースをWEBを通して共有する  
事
  - 基礎的なデータの共有
    - 資料データ（6000件以上の書誌）
    - 地点データ（約3000の県）
  - 方言データの共有
    - 約300項目(大多数は語彙)について各々およそ  
1000地点調査
    - 方言（解釈）地図の共有
    - 地図集編纂のため

## システムの構成

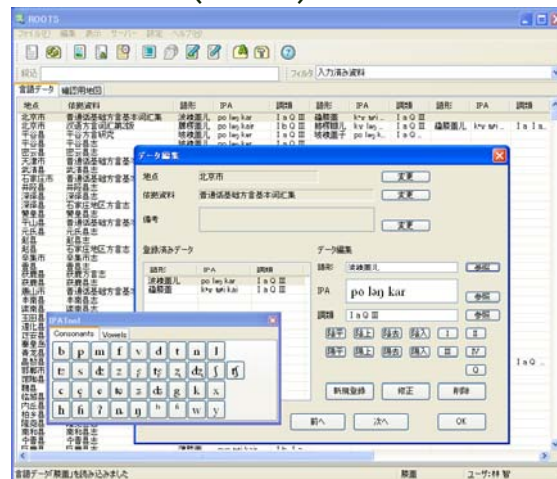
- サーバ側
  - Linuxサーバ
  - WEBサーバ(Apache+PHP)
  - アプリケーションサーバ(Tomcat)
  - データベース(MySQL等)
- クライアント側
  - 言語データ編集ソフト(ROOTS)
  - 地図編集ソフト(Wonderland)
- XMLでサーバ・クライアント間のデータ授受

## 参照関係



## 言語データ

- 各項目について語形、音声、声調などを地点ごとに入力
- 専用クライアント(ROOTS)を用いて全国各地で入力作業



## 方言データのXML表現

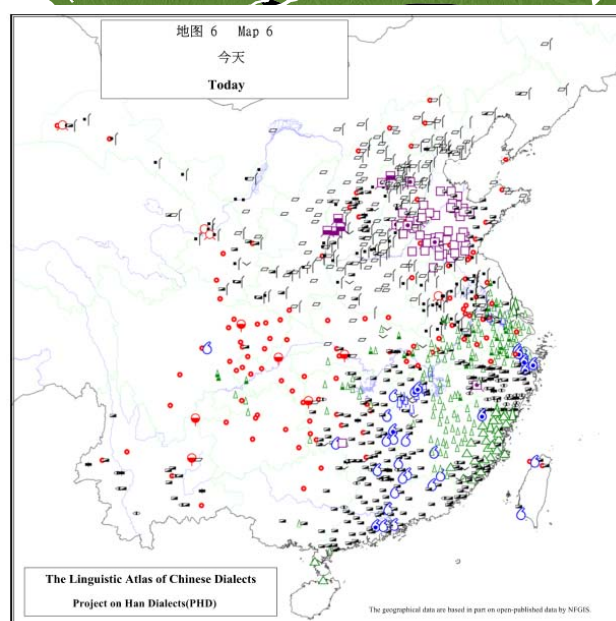
- `<Record>`
- `<Reference>002320</Reference>`
- `<Locality>X0010700</Locality>`
- `<Modifier>Yantian</Modifier>`
- `<DataSet>`
  - `<Artic>IV II III</Artic>`
  - `<IPA>kʰ tɿŋ kər</IPA>`
  - `<Hanzi>圪顶盖儿</Hanzi>`
- `</DataSet>`
- `<DataSet>`
  - `<Artic>IV Q III</Artic>`
  - `<IPA>kʰ lɿŋ kər</IPA>`
  - `<Hanzi>圪灵盖儿</Hanzi>`
- `</DataSet>`
- `</Record>`

- 各Record要素にメタ情報と、複数のDataSet要素がふくまれる。
- DataSet要素には、Hanzi/IPA/Artic要素が含まれる。

## 方言地図の作成

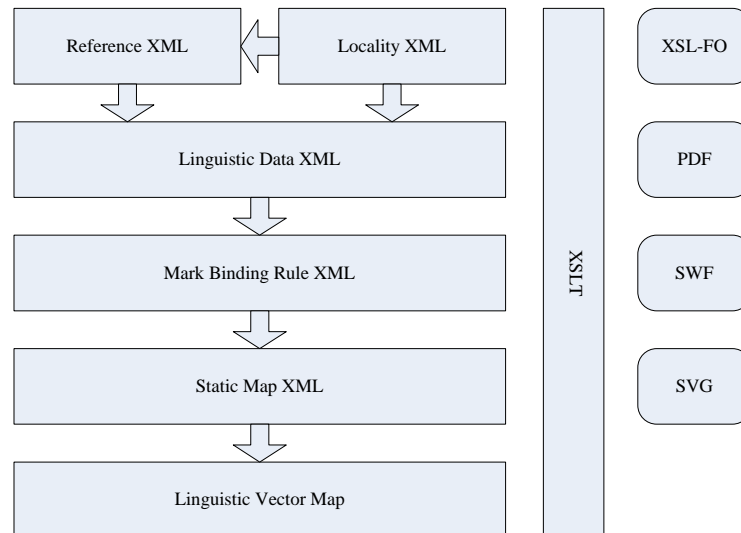
- WEB上の方言データベースから作成  
 複数項目の言語データを同時に読み込める所が画期的
- 読み込んだ方言データをルールに基づきグループ化 → 言語学的な洞察力とArtistic senseが必要
  - 文字列一致 (正規表現)
  - 論理演算 (AND/OR/NOT)
  - グループにマーク(地図上に表示される記号)の割り当て
- これにより様々な方言地図を作成。それらに言語学的な解釈を加える。

## 方言地図の作成例





## Data Flow of the PHD System



## 新PHD以降の主な成果(2004-)

- 方言データの入力
  - ▣ 313項目について約100人の入力者(多くは研究協力者)で入力
- 方言(解釈)地図の作成
  - ▣ 514件の方言地図の作成
- 地図集の発行
  - ▣ 科研報告書地図集 2006年3月, 2007年3月
  - ▣ 『漢語方言解釈地図集』 2009年12月(予定)  
 地図+解説(中国語及び英語で執筆)  
 科研費出版助成を申請中